PATENT
Attorney Docket No. SRC027
Client No. 80404.0032
Express Mail No. EV335405548US

SWITCH/NETWORK ADAPTER PORT INCORPORATING SHARED MEMORY RESOURCES SELECTIVELY ACCESSIBLE BY A DIRECT EXECUTION LOGIC ELEMENT AND ONE OR MORE DENSE LOGIC DEVICES

CROSS REFERENCE TO RELATED PATENT APPLICATIONS

The present invention is a continuation-in-part application and is related to, and claims priority from, United States Patent Application Ser. No. 10/340,390 filed January 10, 2003 for: "Switch/Network Adapter Port Coupling a Reconfigurable Processing Element to One or More Microprocessors for Use With Interleaved Memory Controllers, which is a continuation-in-part application and is related to, and claims priority from, United States Patent Application Ser. No. 09/932,330 filed August 17, 2001 for: "Switch/Network Adapter Port for Clustered Computers Employing a Chain of Multi-Adaptive Processors in a Dual In-Line Memory Module Format" which is a continuation-in-part of Patent Application Serial No. 09/755,744 filed January 5, 2001 which is a divisional patent application of United States Patent Application Serial No. 09/481,902 filed January 12, 2000, now United States Patent No. 6,247,110, which is a continuation of Serial No. 08/992,763, filed December 17, 1997, now United States Patent No. 6,076,152, all of which are assigned to SRC Computers, Inc., Colorado Springs, Colorado, the assignee of the present invention, the disclosures of which are herein specifically incorporated in their entirety by this reference.

BACKGROUND OF THE INVENTION

The present invention relates, in general, to the field of reconfigurable processor-based computing systems. More particularly, the present invention relates to a switch/network adapter port incorporating shared memory resources selectively accessible by a direct execution logic element (such as a reconfigurable computing element comprising one or more field programmable gate arrays "FPGAs") and one or more dense logic devices comprising commercially available microprocessors, digital signal processors ("DSPs"), application specific integrated circuits ("ASICs") and other typically fixed logic components having relatively high clock rates.

As disclosed in one or more representative embodiments illustrated and described in the aforementioned patents and patent applications, SRC Computers, Inc. proprietary Switch/Network Adapter Port technology (SNAP™, a trademark of SRC Computers, Inc., assignee of the present invention) has previously been enhanced such that the signals from two or more dual in-line memory module ("DIMM") (or Rambus™ in-line memory module "RIMM") slots are routed to a common control chip.

Physically, in a by-two configuration, two DIMM form factor switch/network adapter port boards may be coupled together using rigid flex circuit construction to form a single assembly. One of the DIMM boards may also be populated with a control field programmable gate array ("FPGA") which may have the signals from both DIMM slots routed to it. The control chip then samples the data off of both slots using the independent clocks of the slots. The data from both slots is then used to form a data packet that is then

sent to other parts of the system. In a similar manner, the technique may be utilized in conjunction with more than two DIMM slots, for example, four DIMM slots in a four-way interleaved system.

In operation, an interleaved memory system may use two or more memory channels running in lock-step wherein a connection is made to one of the DIMM slots and the signals derived are used in conjunction with the original set of switch/network adapter port board signals. In operation, this effectively doubles (or more) the width of the data bus into and out of the memory. This technique can be implemented in conjunction with the proper selection of a memory and input/output ("I/O") controller ("North Bridge") chip that supports interleaved memory.

Currently described in the literature is a reconfigurable computing development environment called "Pilchard" which plugs into a personal computer DIMM slot. See, for example, "Pilchard - A Reconfigurable Computing Platform with Memory Slot Interface" developed at the Chinese University of Hong Kong under a then existing license and utilizing SRC Computers, Inc. technology. The Pilchard system, and other present day systems rely on relatively long column address strobe ("CAS") latencies to enable the FPGA to process the memory transactions and are essentially slaves to the memory and I/O controller.

With the speed gap ever increasing between the processor speeds and the memory subsystem, processor design has been optimized to keep the cache subsystem filled with data that will be needed by the program currently executing on the processor. Thus, the processor itself is becoming less efficient at performing the large block transfers that may be

required in certain systems utilizing currently available switch/network devices.

SUMMARY OF THE INVENTION

In order to increase processor operational efficiency in conjunction with a switch/network adapter port, the present invention advantageously incorporates and properly allocates memory resources, such as dynamic random access memory ("DRAM"), located on the module itself. Functionally, this memory appears to the dense logic device (e.g. a microprocessor) to be like other system memory and no time penalties are incurred when reading to, or writing from, it.

Through the use of an access coordination mechanism, the control of this memory can be handed off to the switch/network adapter port memory controller. Once in control, the controller can move data between the memory resources and the computer network, based for example, on control parameters that may be located in on-board registers. This data movement is performed at the maximum rate that the memory devices themselves can sustain, thereby providing the highest performance link to the other network devices such as direct execution logic devices such as Multi-Adaptive Processing elements (MAP® a trademark of SRC Computers, Inc.), common memory boards and the like.

Unlike the Pilchard system described previously, the system and method of the present invention does not need to rely on relatively long CAS memory latencies to enable the associated FPGA to process the memory transactions. Moreover, the system and method of the present invention functions as a true peer to

the system memory and I/O controller and access to the shared memory resources is arbitrated for between the memory and I/O controller and the switch/network adapter port controller.

Further, with increasing system security demands, as well as other functions that require unique memory address access patterns, the addition of a programmable memory controller to the system/network adapter port control unit enables this improved system to meet these needs. Functionally, the memory controller is enabled such that the address access patterns utilized in the performance of the data movement to and from the collocated memory resources is programmable. This serves to effectively eliminate the performance penalty that is common when performing scatter/gather and other similar functions.

In a representative embodiment of the present invention disclosed herein, the memory and I/O controller, as well as the enhanced switch/network adapter port memory ("SNAPM™") controller, can control the common memory resources on the SNAPM modules through the inclusion of various data and address switches (e.g. field effect transistors "FETs", or the like) and tri-stable latches. These switching resources and latches are configured such that the data and address lines may be driven by either the memory and I/O controller or the SNAPM memory controller while complete DIMM (and RIMM or other memory module format) functionality is maintained. Specifically, this may be implemented in various ways including the inclusion of a number of control registers added to the address space accessible by the memory and I/O controller which are used to coordinate the use of the shared memory resources.

In operation, when the memory and I/O controller is in control, the SNAPM memory controller is barred from accessing the DRAM memory. Conversely, when the SNAPM memory controller is in control, the address/control and data buses from the memory and I/O controller are disconnected from the DRAM memory. However, the SNAPM memory controller continues to monitor the address and control bus for time critical commands such as memory refresh commands. Should the memory and I/O controller issue a refresh command while the SNAPM memory controller is in control of the DRAM memory, it will interleave the refresh command into its normal command sequence to the DRAM devices. Additionally, when the memory and I/O controller is in control, the SNAPM modules monitor the address and command bus for accesses to any control registers located on the module and can accept or drive replies to these commands without switching control of the collocated memory resources.

Functionally, the SNAPM controller contains a programmable direct memory access ("DMA") engine which can perform random access and other DMA operations based on the state of any control registers or in accordance with other programmable information. The SNAPM controller is also capable of performing data re-ordering functions wherein the contents of the DRAM memory can be read out and then rewritten in a different sequence.

Particularly disclosed herein is a computer system comprising at least one dense logic device, a controller for coupling the dense logic device to a control block and a memory bus, a plurality of memory module slots coupled to the memory bus, an adapter port including shared memory resources associated with

a subset of the plurality of memory module slots and a direct execution logic element coupled to the adapter port. The dense logic device and the direct execution logic element may both access the shared memory resources. In a preferred embodiment, the adapter port may be conveniently provided in a DIMM, RIMM or other memory module form factor.

Also disclosed herein is a computer system comprising at least one dense logic device, an interleaved controller for coupling the dense logic device to a control block and a memory bus, a plurality of memory slots coupled to the memory bus, an adapter port including shared memory resources associated with at least two of the memory slots and a direct execution logic element coupled to at least one of the adapter ports.

Further disclosed herein is a computer system including an adapter port for electrical coupling between a memory bus of the computer system and a network interface. The computer system comprises at least one dense logic device coupled to the memory bus and the adapter port comprises a memory resource associated with the adapter port and a control block for selectively enabling access by the dense logic device to the memory resource. In a particular embodiment disclosed herein, the computer system may further comprise an additional adapter port having an additional memory resource associated with it and the control block being further operative to selectively enable access by the dense logic device to the additional memory resource.

Broadly, the system and method of the present invention disclosed herein includes a switch/network adapter port with collocated memory that may be

isolated to allow peer access to the memory by either a system memory and I/O controller or switch/network adapter port memory controller. The switch/network adapter port with on-board memory disclosed may be utilized as an interface itself and also allows the switch/network adapter port memory controller to operate directly on data retained in the shared memory resources. This enables it to prepare the data for transmission in operations requiring access to a large block of non-sequential data, such as scatter and gather. The system and method of the present invention described herein further discloses a switch/network adapter port with shared memory resources which incorporates a smart, fully parameterized DMA engine providing the capability of performing scatter/gather and other similar functions.

BRIEF DESCRIPTION OF THE DRAWINGS

The aforementioned and other features and objects of the present invention and the manner of attaining them will become more apparent and the invention itself will be best understood by reference to the following description of a preferred embodiment taken in conjunction with the accompanying drawings, wherein:

Fig. 1 is a functional block diagram of a switch/network adapter port for a clustered computing system employing a chain of multi-adaptive processors in a DIMM format functioning as direct execution logic to significantly enhance data transfer rates over that otherwise available from the peripheral component interconnect ("PCI") bus;

Fig. 2A is a functional block diagram of an exemplary embodiment of a switch/network adapter port

incorporating collocated shared memory resources illustrating in a by-two configuration of interleaved DIMM slot form factor SNAPM elements coupled to a common SNAPM memory control element for coupling to a cluster interconnect fabric including one or more direct execution logic devices such as MAP® elements;

Fig. 2B is a further functional block diagram of another exemplary embodiment of a switch/network adapter port incorporating collocated shared memory resources in accordance with the present invention illustrating a by-four configuration of interleaved DIMM slot form factor SNAPM elements coupled to a common SNAPM memory control element;

Fig. 3 is a functional block diagram of a representative embodiment of a by-two SNAPM system in accordance with the present invention comprising a pair of circuit boards, each of which may be physically and electrically coupled into one of two DIMM memory slots, and one of which may contain a SNAPM control block in the form of a field programmable gate array ("FPGA") functioning as the SNAPM memory control block of the preceding Figs. 2A and 2B;

Fig. 4A is a corresponding functional block diagram of the embodiment of the preceding figure wherein the memory and I/O controller drives the address/control and data buses for access to the shared memory resources of the SNAPM elements through the respective address and data switches; and

Fig. 4B is an accompanying functional block diagram of the embodiment of Fig. 3 wherein the SNAPM memory control block provides access to the shared memory resources and disconnects the address/control

and data buses from the system memory and I/O controller.

DESCRIPTION OF A REPRESENTATIVE EMBODIMENT

With reference now to Fig. 1, a functional block diagram of an exemplary embodiment of a computer system 100 is shown comprising a switch/network adapter port for clustered computers employing a chain of multi-adaptive processors functioning as direct execution logic elements in a DIMM format to significantly enhance data transfer rates over that otherwise available from the peripheral component interconnect ("PCI") bus.

In the particular embodiment illustrated, the computer system 100 includes one or more dense logic devices in the form of processors 102_0 and 102_1 which are coupled to an associated memory and I/O controller 104 (e.g. a "North Bridge"). In the operation of the particular embodiment illustrated, the controller 104 sends and receives control information from a separate PCI control block 106. It should be noted, however, that in alternative implementations of the present invention, the controller 104 and/or the PCI control block 106 (or equivalent) may be integrated within the processors 102 themselves and that the control block 106 may also be an accelerated graphics port ("AGP") or system maintenance ("SM") control block. The PCI control block 106 is coupled to one or more PCI card slots 108 by means of a relatively low bandwidth PCI bus 110 which allows data transfers at a rate of substantially 256 MB/sec. In alternative embodiments, the card slots 108 may alternatively comprise PCI-X, PCI Express, accelerated graphics port ("AGP") or system maintenance ("SM") bus connections.

The controller 104 is also conventionally coupled to a number of DIMM slots 114 by means of a much higher bandwidth DIMM bus 116 capable of data transfer rates of substantially 2.1 GB/sec. or greater. In accordance with a particular implementation of the system shown, a DIMM MAP® element 112 may be associated with, or physically located within, one of the DIMM slots 114. Control information to or from the DIMM MAP® element 112 may be provided by means of a connection 118 interconnecting the PCI bus 110 and the DIMM MAP® element 112. The DIMM MAP® element 112 then may be coupled to another clustered computer MAP® element by means of a cluster interconnect fabric connection 120 connected to MAP® chain ports. It should be noted that, the DIMM MAP® element 12 may also comprise a Rambus™ DIMM ("RIMM") MAP® element.

Since the DIMM memory located within the DIMM slots 114 comprises the primary storage location for the microprocessor(s) 102_0, 102_1, it is designed to be electrically very "close" to the processor bus and thus exhibit very low latency. As noted previously, it is not uncommon for the latency associated with the DIMM to be on the order of only 25% of that of the PCI bus 110. By, in essence, harnessing this bandwidth as an interconnect between computer systems 100, greatly increased cluster performance may be realized as disclosed in the aforementioned patents and patent applications.

To this end, by placing the DIMM MAP® element 112 in one of the PC's DIMM slots 114, its control chip can accept the normal memory "read" and "write" transactions and convert them to a format used by an interconnect switch or network. To this end, each MAP® element 112 may also include chain ports to

enable it to be coupled to other MAP® elements 112. Through the utilization of the chain port to connect to the external clustering fabric over connection 120, data packets can then be sent to remote nodes where they can be received by an identical board. In this particular application, the DIMM MAP® element 112 would extract the data from the packet and store it until needed by the receiving processor 102.

This technique results in the provision of data transfer rates several times higher than that of any currently available PC interface such as the PCI bus 110. However, the electrical protocol of the DIMMs is such that once the data arrives at the receiver, there is no way for a DIMM module within the DIMM slots 114 to signal the microprocessor 102 that it has arrived, and without this capability, the efforts of the processors 102 would have to be synchronized through the use of a continued polling of the DIMM MAP® elements 112 to determine if data has arrived. Such a technique would totally consume the microprocessor 102 and much of its bus bandwidth thus stalling all other bus agents.

To avoid this situation, the DIMM MAP® element 112 may be further provided with the connection 118 to allow it to communicate with the existing PCI bus 110 which could then generate communications packets and send them via the PCI bus 110 to the processor 102. Since these packets would account for but a very small percentage of the total data moved, the low bandwidth effects of the PCI bus 110 are minimized and conventional PCI interrupt signals could also be utilized to inform the processor 102 that data has arrived. In accordance with another possible implementation, the system maintenance ("SM") bus (not

shown) could also be used to signal the processor 102. The SM bus is a serial current mode bus that conventionally allows various devices on the processor board to interrupt the processor 102. In an alternative embodiment, the accelerated graphics port ("AGP") may also be utilized to signal the processor 102.

With a DIMM MAP® element 112 associated with what might be an entire DIMM slot 114, the system will allocate a large block of addresses, typically on the order of 1 GB, for use by the DIMM MAP® element 112. While some of these can be decoded as commands, many can still be used as storage. By having at least as many address locations as the normal input/output ("I/O") block size used to transfer data from peripherals, the conventional Intel™ chip sets used in most personal computers (including controller 104) will allow direct I/O transfers into the DIMM MAP® element 112. This then allows data to arrive from, for example, a disk and to pass directly into a DIMM MAP® element 112. It then may be altered in any fashion desired, packetized and transmitted to a remote node over connection 120. Because both the disk's PCI bus 110 and the DIMM MAP® element 112 and DIMM slots 114 are controlled by the PC memory controller 104, no processor bus bandwidth is consumed by this transfer.

It should also be noted that in certain computer systems, several DIMMs within the DIMM slots 114 may be interleaved to provide wider memory access capability in order to increase memory bandwidth. In these systems, the previously described technique may also be utilized concurrently in several DIMM slots 114. Nevertheless, regardless of the particular

implementation chosen, the end result is a DIMM-based MAP® element 112 having one or more connections to the PCI bus 110 and an external switch or network over connection 120 which results in many times the performance of a PCI-based connection alone as well as the ability to process data as it passes through the interconnect fabric.

With reference additionally now to Fig. 2A, a functional block diagram of an exemplary embodiment of a switch/network adapter port 200A incorporating collocated common memory resources in accordance with the present invention is shown. In this regard, like structure and functionality to that disclosed with respect to the foregoing figure is here like numbered and the foregoing description thereof shall suffice herefor. The switch/network adapter port with common memory ("SNAPM") 200A is shown in an exemplary by-two configuration of interleaved DIMM slot form factor SNAPM elements 204 (SNAPM A and SNAPM B) each coupled to a common control element 202 (comprising, together with the two SNAPM elements 204 "SNAPM") and with each of the SNAPM elements 204 including respective DRAM memory 206A and 206B in conjunction with associated switches and buses 208A and 208B respectively as will be more fully described hereinafter. In this embodiment, the controller 104 is an interleaved memory controller bi-directionally coupled to the DIMM slots 114 and SNAPM elements 204 by means of a Channel A 216A and a Channel B 216B.

With reference additionally now to Fig. 2B, a functional block diagram of another exemplary embodiment of a switch/network adapter port 200B incorporating collocated common memory resources in accordance with the present invention is shown.

Again, like structure and functionality to that disclosed with respect to the preceding figures is like numbered and the foregoing description thereof shall suffice herefor. The switch/network adapter port 200B with common memory is shown in a by-four configuration of interleaved DIMM slot form factor SNAPM elements 204 coupled to a common SNAPM memory control element 202 (comprising, together with the four SNAPM elements 204 "SNAPM"). In this embodiment, the controller 104 is again an interleaved memory controller bi-directionally coupled to the DIMM slots 114 and SNAPM elements 204 by means of a respective Channel A 216A, Channel B 216B, Channel C 216C and Channel D 216D.

With reference additionally now to Fig. 3, a functional block diagram of a representative embodiment of a by-two SNAPM system 300 in accordance with the present invention is shown. The SNAPM system, in the exemplary embodiment shown, comprises a pair of circuit boards 204, each of which may be physically and electrically coupled into one of two DIMM (RIMM or other memory module form factor) memory slots, and one of which may contain a SNAPM control block 202 in the form of, for example, an FPGA programmed to function as the SNAPM memory control block of the preceding Figs. 2A and 2B.

Each of the SNAPM circuit boards 204 comprises respective collocated common memory resources 206A ("Memory A") and 206B ("Memory B") which may be conveniently provided in the form of DRAM, SRAM or other suitable memory technology types. Each of the memory resources 206A and 206B is respectively associated with additional circuitry 208A and 208B comprising, in pertinent part, respective DIMM

connectors 302A and 302B, a number of address switches 304A and 304B and a number of data switches 306A and 306B along with associated address/control and data buses. The address switches 304A and 304B and data switches 306A and 306B are controlled by a switch direction control signal provided by the SNAPM control block 202 on control line 308 as shown. The address switches 304 and data switches 306 may be conveniently provided as FETs, bipolar transistors or other suitable switching devices. The network connections 120 may be furnished, for example, as a flex connector and corresponds to the cluster interconnect fabric of the preceding figures for coupling to one or more elements of direct execution logic such as MAP® elements available from SRC Computers, Inc.

With reference additionally now to Fig. 4A, a corresponding functional block diagram of the embodiment of the preceding figure is shown wherein the memory and I/O controller (element 104 of Figs. 1, 2A and 2B) drives the address/control and data buses for access to the shared memory resources 206 of the SNAPM elements 204 through the respective address and data switches 304 AND 306 in accordance with the state of the switch direction control signal on control line 308.

With reference additionally now to Fig. 4B, an accompanying functional block diagram of the embodiment of Fig. 3 is shown wherein the SNAPM memory control block 202 provides access to the shared memory resources 206 and disconnects the address/control and data buses from the system memory and I/O controller (element 104 of Figs. 1, 2A and 2B) in accordance with an opposite state of the switch direction control signal on control line 308.

As shown with respect to Figs. 4A and 4B, the memory and I/O controller (element 104 of Fig. 1, 2A and 2B), as well as the SNAPM memory controller 202, can control the common memory resources 206 on the SNAPM modules 204. The switches 304 and 306 are configured such that the data and address lines may be driven by either the memory and I/O controller 104 or the SNAPM memory controller 202 while complete DIMM (and RIMM or other memory module format) functionality is maintained. Specifically, this may be implemented in various ways including the inclusion of a number of control registers added to the address space accessible by the memory and I/O controller 104 which are used to coordinate the use of the shared memory resources 206. In the embodiment illustrated, the least significant bit ("LSB") data lines (07:00) of lines (71:00) and/or selected address bits may be used to control the SNAPM control block 202, and hence, the allocation and use of the shared memory resources 206.

In operation, when the memory and I/O controller 104 is in control, the SNAPM memory controller 202 is barred from accessing the DRAM memory 206 . Conversely, when the SNAPM memory controller 202 is in control, the address/control and data buses from the memory and I/O controller 104 are disconnected from the DRAM memory 206. However, the SNAPM memory controller 202 continues to monitor the address and control bus for time critical commands such as memory refresh commands. Should the memory and I/O controller 104 issue a refresh command while the SNAPM memory controller 202 is in control of the DRAM memory 206, it will interleave the refresh command into its normal command sequence to the DRAM devices. Additionally, when the memory and I/O controller 104

is in control, the SNAPM modules 204 monitor the address and command bus for accesses to any control registers located on the module and can accept or drive replies to these commands without switching control of the collocated memory resources 206.

While there have been described above the principles of the present invention in conjunction with specific module configurations and circuitry, it is to be clearly understood that the foregoing description is made only by way of example and not as a limitation to the scope of the invention. Particularly, it is recognized that the teachings of the foregoing disclosure will suggest other modifications to those persons skilled in the relevant art. Such modifications may involve other features which are already known per se and which may be used instead of or in addition to features already described herein. Although claims have been formulated in this application to particular combinations of features, it should be understood that the scope of the disclosure herein also includes any novel feature or any novel combination of features disclosed either explicitly or implicitly or any generalization or modification thereof which would be apparent to persons skilled in the relevant art, whether or not such relates to the same invention as presently claimed in any claim and whether or not it mitigates any or all of the same technical problems as confronted by the present invention. The applicants hereby reserve the right to formulate new claims to such features and/or combinations of such features during the prosecution of the present application or of any further application derived therefrom.

What is claimed is: